

A Gradient Boosting Machine Learning Model for Predicting Early Mortality in the Emergency Department Triage: Devising a Nine-Point Triage Score



Maximiliano Klug, MD^{1,2}, Yiftach Barash, MD^{1,2}, Sigalit Bechler, M.Sc³,
Yehezkel S. Resheff, Ph.D³, Talia Tron, Ph.D³, Avi Ironi, MD^{2,4}, Shelly Soffer, MD^{1,2},
Eyal Zimlichman, MD^{2,5}, and Eyal Klang, MD^{1,2}

¹Department of Diagnostic Imaging, The Chaim Sheba Medical Center, Ramat Gan, Israel; ²Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; ³Intuit Israel©, Hod Hasharon, Israel; ⁴Emergency Room, The Chaim Sheba Medical Center, Ramat Gan, Israel; ⁵Hospital Management, The Chaim Sheba Medical Center, Ramat Gan, Israel.

BACKGROUND: Emergency departments (ED) are becoming increasingly overwhelmed, increasing poor outcomes. Triage scores aim to optimize the waiting time and prioritize the resource usage. Artificial intelligence (AI) algorithms offer advantages for creating predictive clinical applications.

OBJECTIVE: Evaluate a state-of-the-art machine learning model for predicting mortality at the triage level and, by validating this automatic tool, improve the categorization of patients in the ED.

DESIGN: An institutional review board (IRB) approval was granted for this retrospective study. Information of consecutive adult patients (ages 18–100) admitted at the emergency department (ED) of one hospital were retrieved (January 1, 2012–December 31, 2018). Features included the following: demographics, admission date, arrival mode, referral code, chief complaint, previous ED visits, previous hospitalizations, comorbidities, home medications, vital signs, and Emergency Severity Index (ESI). The following outcomes were evaluated: *early mortality* (up to 2 days post ED registration) and *short-term mortality* (2–30 days post ED registration). A gradient boosting model was trained on data from years 2012–2017 and examined on data from the final year (2018). The area under the curve (AUC) for mortality prediction was used as an outcome metric. Single-variable analysis was conducted to develop a *nine-point triage score* for early mortality.

KEY RESULTS: Overall, 799,522 ED visits were available for analysis. The early and short-term mortality rates were 0.6% and 2.5%, respectively. Models trained on the full set of features yielded an AUC of 0.962 for early mortality and 0.923 for short-term mortality. A model that utilized the nine features with the highest single-variable AUC scores (age, arrival mode, chief complaint, five primary vital signs, and ESI) yielded an AUC of 0.962 for early mortality.

CONCLUSION: The gradient boosting model shows high predictive ability for screening patients at risk of early mortality utilizing data available at the time of triage in the ED.

KEY WORDS: machine learning; gradient boosting; triage; emergency department; early mortality.

J Gen Intern Med 35(1):220–7
DOI: 10.1007/s11606-019-05512-7
© Society of General Internal Medicine 2019

INTRODUCTION

Emergency departments (ED) are becoming increasingly overwhelmed, increasing poor outcomes associated with ED overcrowding.^{1–5} Triage scores^{6,7} aim to optimize the waiting time and prioritize the resource usage according to the severity of the medical condition. The Emergency Severity Index (ESI) is the most widely used triage score.⁸ It is a subjective risk classification of patients, from 1 (most urgent) to 5 (least urgent), based on patients' acuity and resources needed. The ESI relies heavily on provider judgment which can lead to inaccuracy and misclassification.^{9,10} Differentiating between levels 2 and 3 is a challenging task,¹¹ and ESI level 3 is assigned to a largely diverse ill patient group.¹²

Artificial intelligence (AI) algorithms offer advantages for creating predictive clinical applications because of flexibility in handling large datasets from electronic medical records (EMR).¹³ AI algorithms are becoming better at prediction tasks, often outperforming current clinical scoring systems.¹⁴ Several prediction models have been developed using these techniques in the past years trying to improve the triage process.^{10,15–22} A machine learning prediction model may improve the proper identification of patients at greater risk of mortality and would be superior to a nonsystematic experience-based assessment.^{10,13}

In this study, we aimed to evaluate a state-of-the-art machine learning model for predicting mortality at the triage level. By validating this automatic tool, our purpose

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-019-05512-7>) contains supplementary material, which is available to authorized users.

Received December 21, 2018

Revised July 12, 2019

Accepted October 1, 2019

Published online November 1, 2019

is to improve the categorization of patients in the ED using readily available data from the EMR at the time of arrival.

DESIGN: MATERIALS AND METHODS

An institutional review board (IRB) approval was granted for this retrospective study.

Study Cohort

Information of consecutive adult patients (ages 18–100) admitted at the academic ED of one tertiary center were retrieved from the hospital's EMR. This acute care hospital has approximately 1700 beds and gets 185,000 total emergency department visits per year. Universal health coverage is provided to everyone as stipulated by the public healthcare system.

The study time frame was from January 1, 2012, to December 31, 2018.

Data Acquisition

Data retrieved were information available at the triage level:

Demographics: age, sex

Admission date: retrieved as a timestamp variable

Arrival mode: either walk-in, by ambulance (BLS), or by intensive care ambulance (ALS)

Referral code: either independent or referred by a physician

Chief complaint: which was recorded in two methods: (1) a structured list of 122 chief complaints from a list of the Israeli Health Ministry guidelines available at the ED, (2) a two-word free text chief complaint obtained by a triage nurse

Previous ED visits: dates of all previous visits to our ED during the study time frame

Previous hospitalizations: dates of all previous hospitalizations in our hospital during the study time frame

Comorbidities: coded as International Classification of Diseases (ICD9) records

Home medications: coded using World Health Organization (WHO) Anatomical Therapeutic Chemical Classification System (ATC)

Vital signs: temperature (T°), heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), oxygen saturation (SO_2)

ESI score

Mortality dates were obtained from the EMR and from the Ministry of internal mortality records (when the death occurred outside the hospital), and the duration in days from admission to overall mortality was computed. The following features were used as patient outcome endpoints: early mortality, defined as mortality up to 2 days from registration to the ED, and short-term mortality, defined as mortality 2–30 days from registration to the ED.

Data Pre-processing

Data Cleaning. Patients aged 18–100 years were included. Records of patients with erroneous unreasonable values were removed: vital signs were limited to systolic blood pressure (SBP) < 300 mmHg, diastolic blood pressure (DBP) < 250 mmHg, pulse < 300 beats/min, temperature ranging between 25 and 45 °C, oxygen saturation \leq 100%.

Data Encoding. Input features: all categorical factors were encoded as numerical. For comorbidities, home medications, and unstructured chief complaint high cardinality variables, we used target encoding for embedding. For home medications, we also used one hot encoding using ATC pharmacologic subgroups. Using the lists of previous ED visits and previous hospitalizations at our hospital, we compiled the following: number of previous ED visits, number of previous hospitalizations, number of days to most recent previous ED visit, number of days to most recent previous hospitalization. From the current ED visit timestamp we retrieved the year, month, day, and hour.

Output outcome was encoded as a binary variable (1 = patient died in a selected time frame, 0 = patient did not die in a selected time frame).

Machine Learning Model

The algorithms were programmed using Python (Version 3.6.5 64bits) and the XGBoost open-source library (version 0.8) with the scikit-learn wrapper (version 0.19.2). We addressed class imbalance between mortality and non-mortality cases by using the XGBoost class weights scaling feature. The training data (X , with multiple features) was configured as input to predict a target variable (Y) which was considered a label. Statistical calculations were performed using Python. Computations were done on an Intel I7 CPU and NVIDIA GeForce GTX 1080Ti GPU computer.

Gradient boosting²³ is a machine learning algorithm where multiple weak learners (tree based classifiers) are trained to augment each other and produce superior results. It differs from random forests (RF)^{24,25} in which trees are learned sequentially and based on the performance of all previous trees. In gradient boosting, at each stage, a new decision tree is learned with the aim to correct errors made by existing trees. As a non-linear method, it naturally outperforms linear models²⁶ when higher order relationships exist in the data. Gradient boosting has also surpassed other machine learning algorithms in a number of data challenges.^{23,27} Recent works described its potential in the medical field.^{26–33}

Analysis

Continuous features are reported as the median with the spread reported as interquartile range (IQR). Categorical elements are reported as percentages. The model's performance was assessed using the area under the curve (AUC) metric. We have calculated AUC both for training (years 2012–2017) and

Table 1 Study Cohort Characteristics (n = 799,522)

	Entire cohort (n = 799,522)	Early mortality (n = 4561, 0.6%)	Short-term mortality (n = 19,647, 2.5%)
Age, median (IQR) (y)	55.0 (35.0–72.0)	81.0 (69.0–88.0)	78.0 (66.0–87.0)
Female (N (%))	38,7495 (48.5)	2220 (48.7)	9104 (46.3)
Mode of arrival			
Walk-in (N (%))	601,972 (75.3)	674 (14.8)	7567 (38.5)
Ambulance (BLS) (N (%))	147,963 (18.5)	2798 (61.3)	8490 (43.2)
Intensive care ambulance (ALS) (N (%))	8755 (1.1)	145 (3.2)	532 (2.7)
Other (N (%))	5762 (0.7)	297 (6.5)	455 (2.3)
ESI (1–5)			
1:	4630 (0.6%)	956 (21.0%)	752 (3.8%)
2:	26,723 (3.3%)	942 (20.7%)	2,199 (11.2%)
3:	506,014 (63.3%)	1900 (41.7%)	12,946 (65.9%)
4:	149,778 (18.7%)	57 (1.2%)	718 (3.7%)
5:	25,623 (3.2%)	9 (0.2%)	51 (0.3%)
SBP, median (IQR) (mmHg)	129.0 (116.0–147.0)	110.0 (89.0–135.0)	122.0 (105.0–141.0)
DBP, median (IQR) (mmHg)	75.0 (67.0–83.0)	64.0 (50.0–78.0)	69.0 (59.0–79.0)
Heart rate, median (IQR) (b/min)	80.0 (70.0–92.0)	93.0 (75.0–112.0)	89.0 (75.0–105.0)
Temperature, median (IQR) (Celsius)	36.8 (36.6–37.0)	36.7 (36.3–37.4)	36.8 (36.5–37.2)
Saturation, median (IQR) (SaO ₂)	98.0 (96.0–99.0)	93.0 (86.0–97.0)	95.0 (92.0–98.0)
Number of home medications, median (IQR) (N)	0.0 (0.0–4.0)	6.0 (0.0–9.0)	6.0 (1.0–10.0)
Number of previous ED visits, median (IQR) (N)	0.0 (0.0–2.0)	1.0 (0.0–3.0)	2.0 (0.0–4.0)
Number of previous hospitalizations, median (IQR) (N)	0.0 (0.0–1.0)	1.0 (0.0–2.0)	1.0 (0.0–2.0)

IQR, interquartile range; ESI, Emergency Severity Index; y, year; SD, standard deviation; SBP, systolic blood pressure; DBP, diastolic blood pressure; ED, emergency department; SaO₂, oxygen blood saturation; b/min, beats per minute

validation (year 2018). An AUC of 1 indicates perfect outcome prediction in all patients whereas a value of 0.5 means that the model is no better than chance. An AUC value of greater than 0.80 is desirable.¹⁸ Models without including the ESI were also calculated.

Bootstrapping validations (1000 bootstrap resamples) were used to calculate 95% confidence intervals (CI) for AUCs. Youden's index was employed to find the optimal cutoff point on the receiver operating characteristic (ROC) curve in order to calculate sensitivity, specificity, false-positive rate (FPR), negative predictive value (NPV), and positive predictive value (PPV) of the final models. We also evaluated the sensitivity, FPR, NPV, and PPV for fixed specificities of 95% and 97.5%.

Instead of using cross-validation, models were trained on data from the years 2012–2017 and tested on data from the year 2018, ensuring by this way that there was no overfitting and no leakage of information.

Variable Analysis. We evaluated single-variable predictions by using each one as a single input, and early mortality and short-term mortality as targets. These experiments were conducted using the same data split as described before.

We also evaluated variables' importance for both mortality frames by using the XGBoost feature importance property, which shows for each variable how much on average the prediction changes if the variable value changes.

Model Assessment. We used Brier score and calibration plots to evaluate the XGBoost model.

Comparison with Previous Models. We compared the XGBoost model with severity scores Shock Index (SI), Modified Shock Index (MSI), and Aged Shock Index (ASI)^{9,11}

where $SI = HR/SBP$; $MSI = HR/(2/3 \times DBP + 1/3 \times SBP)$; $ASI = Age \times SI$.

Nine-Point Triage Score. The top single most predictive variables examined before were tested as input factors in an early mortality predictive model. The final model was evaluated on the entire cohort, the cohort after dropping of all missing values, and the cohort after dropping only visits with missing structured chief complaint. We compared these models with a logistic regression (LR) model employing the top features as inputs (used one hot encoding for structured chief complaint) evaluated on the data after dropping all cases with missing values.

The top most important features evaluated before were also tested on the entire cohort in an early mortality predictive model.

RESULTS

Study Cohort

A total of 990,864 ED visits were retrospectively retrieved during the 6-year study frame. We excluded 190,609 records for age criteria, and 733 records for vital signs criteria. Thus, 799,522 ED visits which represented 367,219 unique patients were available for analysis. The overall early mortality rate was 4561/799,522 (0.6%) and the short-term mortality rate was 19,647/799,522 (2.5%). Of the 4561 patients with early mortality, 917 (20.1%) died in the ED, 3405 (74.6%) died in hospitalization, 33 (0.7%) died in the internal care unit (ICU), and 206 (4.5%) died after being discharged from the ED.

Characteristics of the study cohort are presented in Table 1. Table A1 (online) presents the proportions of abnormal range

Table 2 Models’ Performances for Each Mortality Group in the 1-Year Validation Set. Sensitivity, Specificity, FPR, NPV, and PPV Are Calculated for Youden’s Index, and Sensitivity; NPV and PPV Are also Calculated for Fixed Specificities of 95% and 97.5%

	AUC (95% CI)	Cutoff	Sensitivity (95% CI)	Specificity (95% CI)	FPR	NPV (95% CI)	PPV (95% CI)
Early mortality	0.962 (0.956–0.968)	Youden’s index	91.9 (86.1–93.6)	88.2 (87.6–93.6)	1:8.5	99.9 (99.9–100.0)	4.0 (3.7–6.8)
		Fixed specificity 95%	81.2 (78.4–84.1)	95	1:20	99.9 (99.9–99.9)	8.1 (7.5–8.7)
		Fixed specificity 97.5%	67.2 (63.4–70.6)	97.5	1:40	99.8 (99.8–99.8)	12.5 (11.7–13.8)
Short-term mortality	0.923 (0.919–0.926)	Youden’s index	90.5 (88.9–91.8)	80.8 (80.0–82.3)	1:5.2	99.7 (99.6–99.7)	11.8 (11.3–12.6)
		Fixed specificity 95%	54.8 (52.9–56.7)	95	1:20	98.7 (98.6–98.7)	22.4 (22.7–24.5)
		Fixed specificity 97.5%	37.0 (37.3–41.1)	97.5	1:40	98.2 (98.1–98.3)	29.5 (28.3–30.8)

Models’ performance AUC expected cut point ≥ 0.8

CI, confidence interval; AUC, area under the curve; FPR, false-positive rate; NPV, negative predictive value; PPV, positive predictive value

vital signs. The ten most common chief complaints for early and non-early mortality groups are presented in Table A2 (online). The total cumulative mortality up to 30 days, and grouped according to ESI levels, is presented in Figure A3 (online). The proportions of missing data per features are presented in Table A4 (online).

Model Performance

Models were trained using triage variables and tested using 1000 decision trees. The input vector included 308 features. The performances of the models in the validation set for both mortality frames are reported in Table 2. Models yielded an AUC of 0.962 for early mortality and an AUC of 0.923 for short-term mortality. Mortality rates for training and validation sets were represented separately in Figure 1. Models without including the ESI are shown in Table A5 (online).

Brier scores for the early mortality and short-term mortality models were 0.004 and 0.022, respectively. Figure A6 (online) present the calibration curves.

Variable Analysis

Table 3 presents the top ten highest single-variable AUCs for each mortality group. For early mortality, age, arrival mode, and structured chief complaint had the highest AUCs (0.810, 0.809, and 0.787, respectively). For short-term mortality age, comorbidities coded by ICD9 and structured chief complaint had the highest AUCs (0.767, 0.754, and 0.749).

Figures A7 and A8 (online) show the top ten features calculated by the XGBoost feature importance property for early mortality and short-term mortality respectively.

Comparison Between Gradient Boosting Model and Previous Models

Table 4 presents comparisons between the AUCs of the gradient boosting model, SI, MSI, and ASI. From the evaluated previous models, ASI showed the highest predictive ability with an AUC of 0.858 for early mortality and 0.834 for short-term mortality.

Figure 2 presents histograms of SBP, HR, SI, and ASI of patients who died and did not die within 2 days of ED admission. The histogram of ASI shows the greatest differentiation between the two groups.

Nine-Point Triage Score

The top ten single predictive variables were experimented as factors in an early mortality predictive model. Number of comorbidities was dropped as it requires previous history of the patient and we wanted to evaluate a model that does not require such data. Consequently, we assessed a gradient boosting model for early mortality prediction which included only nine elements: age, arrival mode, structured chief complaint, vital signs (T°, SO₂, HR, SBP, and DBP), and ESI. Table 5 presents the AUCs, sensitivities, and specificities of the XGBoost model using the selected nine variables as inputs and early mortality as output. These experiments were conducted on three cohorts: the entire dataset, the cohort after dropping all cases with missing data, the cohort after dropping only cases with missing structured chief complaint. A LR model after dropping all cases with missing data was also added for comparison.

An early mortality prediction model using the top ten features from XGBoost feature importance analysis was also calculated, with its properties presented in Table A9 (online).

DISCUSSION

The main aim of this project was to improve the categorization of patients in the ED by constructing a model that predicts mortality outcomes at the triage level. A prediction model can integrate all available information and facilitate the identification of patients with a higher mortality risk who might be missed. As a tool for health care providers in the decision-making process, it may be used to ensure rapid treatment, and flag high-risk patients subjectively under-triaged.^{10,13}

The non-linear gradient boosting model demonstrated a high predictive ability with an AUC of 0.962 AUC for early

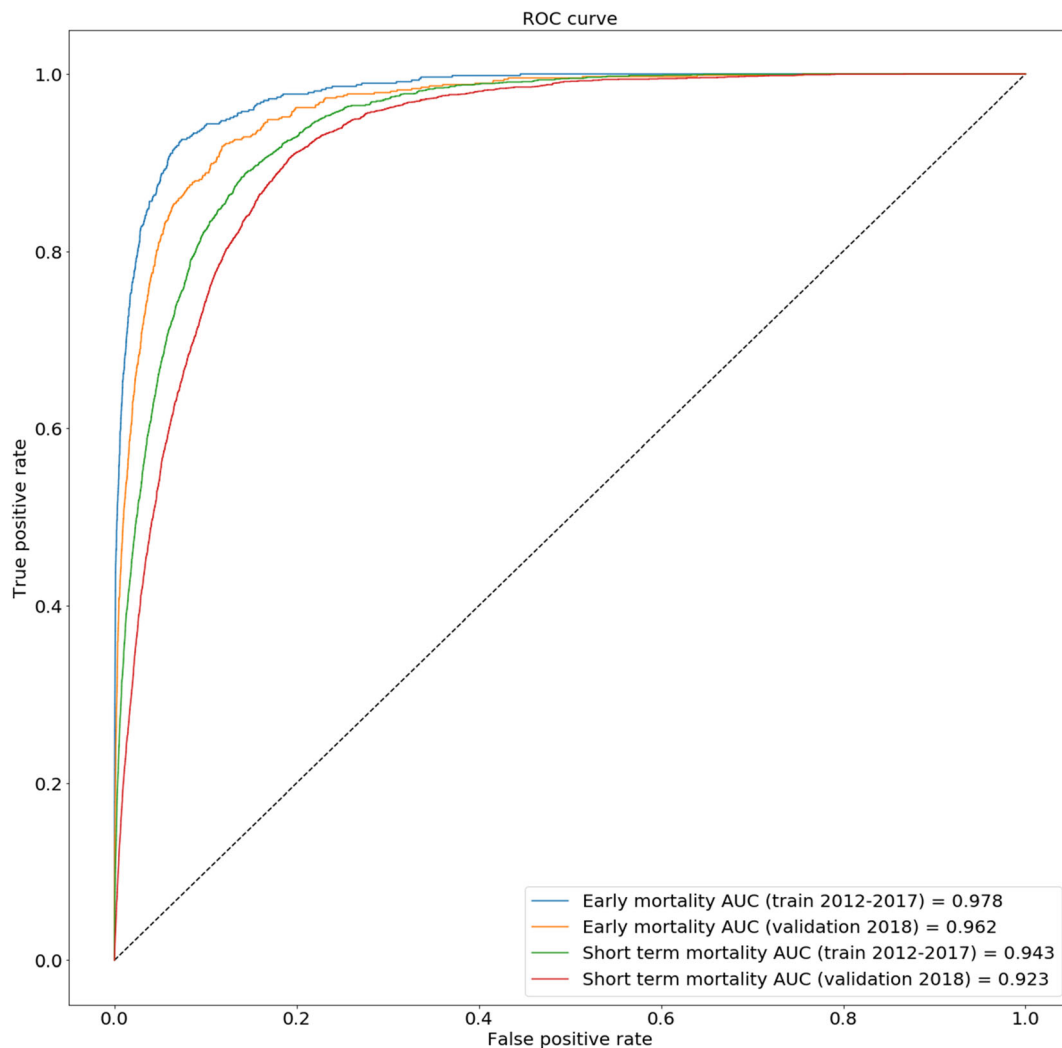


Fig. 1 Receiver operating characteristic (ROC) curves comparing models' performances for each mortality group.

mortality using only triage available features, and a decreased ability for short-term mortality. In our opinion, a possible explanation for this difference could be that the severity of physiologic abnormalities at initial presentation is not as much

related with short-term mortality compared with early mortality.

When analyzing single variables, age and structured chief complaint are the highest predictors of mortality for all time frames. Other factors that foretell early mortality are those reflecting the acute condition of the patient, such as arrival mode, ESI, and vital signs. For short-term mortality, characteristics reflecting the patient's background, such as comorbidities and home medications, seem to be better predictors.

Table 3 Highest Single AUCs for Each Mortality Group for the Ten Clinical Features Calculated by Single-Variable Analysis

Early mortality		Short-term mortality	
Age	0.810	Age	0.767
Arrival mode	0.809	Comorbidities ICD9	0.754
Structured chief complaint	0.787	Structured chief complaint	0.749
Oxygen saturation	0.783	Number of comorbidities	0.746
ESI	0.762	Type of home medications	0.737
Temperature	0.733	Number of home medications taken	0.734
Heart rate	0.728	Oxygen saturation	0.728
Diastolic blood pressure	0.725	Unstructured chief complaint	0.708
Number of comorbidities	0.724	Number of days to previous ED visit	0.677
Systolic blood pressure	0.723	Arrival mode	0.656

AUC, area under the Curve; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; ESI, Emergency Severity Index

Table 4 Mortality Prediction Comparative Analysis (AUCs)

	Early mortality	Short-term mortality
Gradient boosting model	0.962 (95%CI, 0.956–0.968)	0.923 (95%CI, 0.919–0.926)
SI	0.742 (95%CI, 0.721–0.765)	0.664 (95%CI, 0.654–0.675)
MSI	0.751 (95%CI, 0.730–0.772)	0.686 (95%CI, 0.675–0.697)
ASI	0.858 (95%CI, 0.841–0.874)	0.834 (95%CI, 0.827–0.841)

AUC, area under the curve; CI, confidence interval; ESI, Emergency Severity Index; SI, Shock Index; MSI, Modified Shock Index; ASI, Age Shock Index

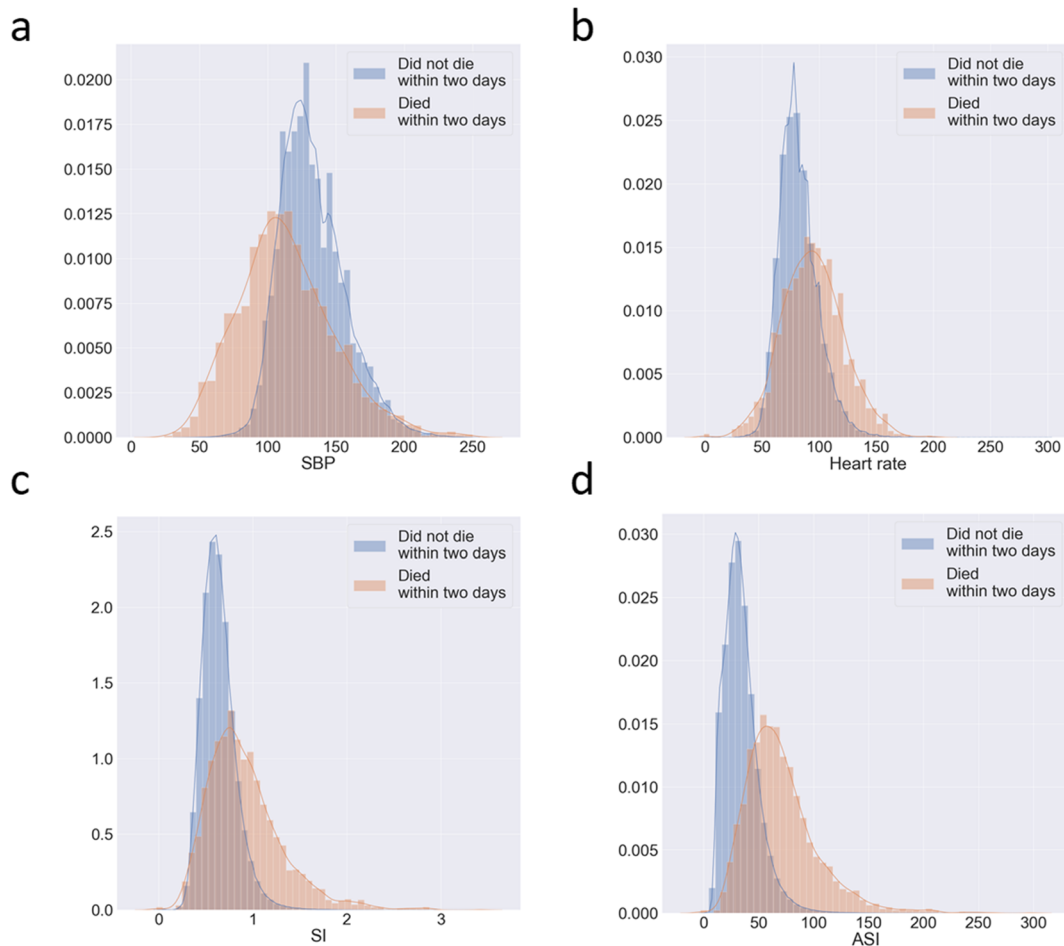


Fig. 2 Frequency histograms of the recorded SBP (a), HR (b), SI (c), and ASI (d) for subject who died and did not die within 2 days of ED admission. *SBP*, systolic blood pressure. *HR*, heart rate. *SI*, Shock Index. *ASI*, Age Shock Index.

To date, the most widely used triage scoring tool is the ESI score.⁸ This score is a subjectively assessed five-level ED triage algorithm that provides risk stratification of patients, from 1 (most urgent) to 5 (least urgent), based on patients' acuity and resources needed.

Several previous studies have developed scores for mortality prediction as a way to improve triage classification, and some of them used AI algorithms.^{9–11,15–21} Levin et al. utilized a random forest E-triage prediction model that had AUCs ranging from 0.90 to 0.92 for critical care outcome (in-

hospital mortality or direct admission to an intensive care unit).¹⁰ A major strength of our study was the significantly larger number of ED visits in comparison with previous studies, which could have increased the model's performance. When examining previous models, the ASI model⁹ showed impressive results.

Single feature analysis helped us devise the nine-point triage score which included age, arrival mode, structured chief complaint, vital signs (T°, SO₂, HR, SBP, and DBP), and ESI into the gradient boosting model. It showed an AUC of 0.962

Table 5 Mortality Prediction and Statistical Analysis of the XGBoost Model Using Nine-Point Triage Score

	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	FPR	NPV (95% CI)	PPV (95% CI)
XGBoost: entire cohort (<i>n</i> = 799,522; early mortality = 4561)	0.960 (0.954–0.966)	90.1 (0.871–0.920)	89.9 (0.867–0.925)	1:9.9	99.9 (99.9–100.0)	4.6 (3.6–6.0)
XGBoost: cohort after dropping all cases with missing values (<i>n</i> = 344,180; early mortality = 1829)	0.949 (0.939–0.958)	89.1 (82.8–91.6)	87.5 (86.4–92.6)	1:10	99.9 (99.9–100.0)	3.2 (2.8–5.1)
Logistic regression: cohort after dropping all cases with missing values (<i>n</i> = 344,180; early mortality = 1829)	0.937 (0.925–0.948)	86.4 (82.7–90.5)	87.6 (83.6–89.6)	1:8.1	99.9 (99.9–100.0)	3.1 (2.5–3.7)
XGBoost: cohort after dropping cases with missing structured chief complaint (<i>n</i> = 538,258; early mortality = 2989)	0.962 (0.956–0.968)	90.7 (86.7–92.6)	89.6 (88.1–92.8)	1:9.6	99.9 (99.9–100.0)	4.6 (3.9–6.3)

Models' performance AUC expected cut point ≥ 0.8

CI, confidence interval; AUC, area under the curve; FPR, false-positive rate; NPV, negative predictive value; PPV, positive predictive value

for early mortality, similar to the AUC of the full features model. Using a simplified model makes it easier to understand which variables are truly driving the early mortality outcome allowing to improve it.³⁴

Models that use hundreds of variables make manual entry impractical and are difficult to encode within EMR databases. A simplified model with fewer variables would make this task more feasible.³¹

Our study has several limitations: Firstly, it was a retrospective single-institution study, and the sample was homogeneous and may have been subject to local practices which limit its generalizability; also the high performance of ASI casts the doubt if it can outperform the model during external validation. Secondly, we lack information from other institutions about previous ED visits and hospitalizations. Thirdly, do-not-resuscitate (DNR) patients, on-hospice, left-before-being-seen or against-medical-advice could not be excluded. Fourthly, other outcomes like ICU admission or in-hospital mortality were not evaluated.

We believe this study may serve as a proof of concept for the ability to develop AI-based triage prediction models, to be replicated on multi-institute projects by matching/mapping data using a multicenter approach in order to ascertain the predictive accuracy.

The presented model suggests a decision support tool, and does not intend to replace clinical judgment. The interaction with provider intuition can strengthen clinical decision-making, making it more consistent and reducing the risks of over- or under-triage. A strong collaboration between clinicians and machine learning experts is necessary to develop and validate a model which includes the best predicting variables in order to prevent entering large amounts of data without a clinical context.³²

In conclusion, by using data available at the time of triage in the ED, the gradient boosting model showed a high predictive ability while screening patients at risk of early mortality.

Acknowledgments: This research was performed in collaboration with the Intuit data science team as part of the philanthropic framework, *We Care and Give Back*. It was also conducted with the help of ARC - The Innovation Center at Sheba Hospital.

Corresponding Author: Eyal Klang, MD; Department of Diagnostic Imaging The Chaim Sheba Medical Center, Ramat Gan, Israel (e-mail: eyalkla@hotmail.com).

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

- Carter EJ, Pouch SM, Larson EL. The relationship between emergency department crowding and patient outcomes: a systematic review. *J Nurs Scholarsh*. 2014;46(2):106-15.
- Johnson KD, Winkelman C. The effect of emergency department crowding on patient outcomes: a literature review. *Adv Emerg Nurs J*. 2011;33(1):39-54.
- Pines JM, Iyer S, Disbot M, Hollander JE, Shofer FS, Datner EM. The effect of emergency department crowding on patient satisfaction for admitted patients. *Acad Emerg Med*. 2008;15(9):825-31.
- Sun BC, Hsia RY, Weiss RE, Zingmond D, Liang L-J, Han W, et al. Effect of emergency department crowding on outcomes of admitted patients. *Ann Emerg Med*. 2013;61(6):605-11. e6.
- Chiu I-M, Lin Y-R, Syue Y-J, Kung C-T, Wu K-H, Li C-J. The influence of crowding on clinical practice in the emergency department. *Am J Emerg Med*. 2018;36(1):56-60.
- Farrohknia N, Castren M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, et al. Emergency department triage scales and their components: a systematic review of the scientific evidence. *Scand J Trauma Resuscitation Emerg Med*. 2011;19:42.
- Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. *Deutsches Arzteblatt Int*. 2010;107(50):892-8.
- McHugh M, Tanabe P, McClelland M, Khare RK. More Patients Are Triageed Using the Emergency Severity Index Than Any Other Triage Acuity System in the United States. *Acad Emerg Med*. 2012;19(1):106-9.
- Torabi M, Moeinaddini S, Mirafzal A, Rastegari A, Sadeghkhan N. Shock index, modified shock index, and age shock index for prediction of mortality in Emergency Severity Index level 3. *Am J Emerg Med*. 2016;34(11):2079-83.
- Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, et al. Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. *Ann Emerg Med*. 2018;71(5):565-74 e2.
- Torabi M, Mirafzal A, Rastegari A, Sadeghkhan N. Association of triage time Shock Index, Modified Shock Index, and Age Shock Index with mortality in Emergency Severity Index level 2 patients. *Am J Emerg Med*. 2016;34(1):63-8.
- Arya R, Wei G, McCoy JV, Crane J, Ohman-Strickland P, Eisenstein RM. Decreasing Length of Stay in the Emergency Department With a Split Emergency Severity Index 3 Patient Flow Model. *Acad Emerg Med*. 2013;20(11):1171-9.
- Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *npj Digital Med*. 2018;1(1):18.
- Stewart J, Sprivilis P, Dwivedi G. Artificial intelligence and machine learning in emergency medicine. *Emerg Med Australas*. 0(0).
- Coslovsky M, Takala J, Exadaktylos AK, Martinolli L, Merz TM. A clinical prediction model to identify patients at high risk of death in the emergency department. *Intensive Care Med*. 2015;41(6):1029-36.
- Pearl A, Bar-Or R, Bar-Or D. An artificial neural network derived trauma outcome prediction score as an aid to triage for non-clinicians. *Stud Health Technol Inform*. 2008;136:253.
- Dugas AF, Kirsch TD, Toerper M, Korley F, Yenokyan G, France D, et al. An Electronic Emergency Triage System to Improve Patient Distribution by Critical Outcomes. *J Emerg Med*. 2016;50(6):910-8.
- Teubner DJ, Considine J, Hakendorf P, Kim S, Bersten AD. Model to predict inpatient mortality from information gathered at presentation to an emergency department: The Triage Information Mortality Model (TIMM). *Emerg Med Australas*. 2015;27(4):300-6.
- Schuetz P, Hausfater P, Amin D, Haubitz S, Fassler L, Grolimund E, et al. Optimizing triage and hospitalization in adult general medical emergency patients: the triage project. *BMC Emerg Med*. 2013;13:12.
- Barak-Corren Y, Israelit SH, Reis BY. Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. *Emerg Med J*. 2017;34(5):308-14.
- Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. *Acad Emerg Med*. 2011;18(8):844-50.
- Barak-Corren Y, Fine AM, Reis BY. Early Prediction Model of Patient Hospitalization From the Pediatric Emergency Department. *Pediatrics*. 2017;139(5).
- Chen T, Guestrin C, editors. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016: ACM.
- Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
- Biau G, Scornet E. A random forest guided tour. *Test*. 2016;25(2):197-227.
- Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. *PLoS One*. 2018;13(7):e0201016.

27. **Giao Z, Sun N, Li X, Xia E, Zhao S, Qin Y.** Using Machine Learning Approaches for Emergency Room Visit Prediction Based on Electronic Health Record Data. *Stud Health Technol Inform.* 2018;247:111–5.
28. **Goto T, Camargo Jr CA, Faridi MK, Yun BJ, Hasegawa K.** Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. *Am J Emerg Med.* 2018;36(9):1650–4.
29. **Bogle B, Balduino R, Wolk DM, Farag HA, Kethireddy S, Chatterjee A.** et al. Predicting Mortality of Sepsis Patients in a Multi-Site Healthcare System using Supervised Machine Learning. Available at: <https://csce.ucmss.com/cr/books/2018/LFS/CSREA2018/HIM3645.pdf> Accessed July 1, 2019.
30. **Ho EL, Tan I, Lee I, Wu P, Chong H.** Predicting Readmission at Early Hospitalization Using Electronic Health Data: A Customized Model Development. *Int J Integrated Care.* 2017;17(5).
31. **Taylor RA, Moore CL, Cheung K-H, Brandt C.** Predicting urinary tract infections in the emergency department with machine learning. *PLoS One.* 2018;13(3):e0194085.
32. **Hill B, Brown RP, Gabel E, Lee C, Cannesson M, Loohuis LO,** et al. Preoperative predictions of in-hospital mortality using electronic medical record data. *bioRxiv.* 2018:329813.
33. **Maali Y, Perez-Concha O, Coiera E, Roffe D, Day RO, Gallego B.** Predicting 7-day, 30-day and 60-day all-cause unplanned readmission: a case study of a Sydney hospital. *BMC Med Inform Decis Mak.* 2018;18(1):1.
34. **Awad A, Bader-El-Den M, McNicholas J, Briggs J.** Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Inform.* 2017;108:185–95.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.